



Webからの画像データ自動収集と機械学習を用いた オブジェクト識別システムの構築

著者	手塚 太郎
発行年	2013
その他のタイトル	Object Identification System using Web Image Collection and Machine Learning
URL	http://hdl.handle.net/2241/120811

科学研究費助成事業（科学研究費補助金）研究成果報告書

平成 25 年 4 月 25 日現在

機関番号：12102

研究種目：若手研究（B）

研究期間：2009～2012

課題番号：21700121

研究課題名（和文） Webからの画像データ自動収集と機械学習を用いたオブジェクト識別システムの構築

研究課題名（英文） Object Identification System using Web Image Collection and Machine Learning

研究代表者

手塚 太郎（TEZUKA TARO）

筑波大学・図書館情報メディア系・准教授

研究者番号：40423016

研究成果の概要（和文）：

World Wide Web 上から画像を収集し、各単語に対応する画像に共通する画像特徴量とその空間的位置関係を学習させることで、典型的な画像を抽出し、任意の画像に何が写っているかを判別する一般物体認識システムの研究開発を進めた。

研究成果の概要（英文）：

By collecting images from the World Wide Web, a method of extracting visual features and their spatial allocations were developed. Using this, a set of typical images were obtained for each term, and a system of general object recognition was developed.

交付決定額

（金額単位：円）

	直接経費	間接経費	合 計
2009年度	700,000	210,000	910,000
2010年度	500,000	150,000	650,000
2011年度	500,000	150,000	650,000
2012年度	500,000	150,000	650,000
年度			
総 計	2,200,000	660,000	2,860,000

研究分野：情報学

科研費の分科・細目：情報学 メディア情報学・データベース

キーワード：コンテンツ，画像認識

1. 研究開始当初の背景

画像に映っている物体が何であることを認識する一般物体認識は画像認識の研究においてもっとも注目されている課題のひとつであり、同時にまたもっとも困難な課題のひとつとされている。

これは物体の種類が多岐にわたり、またそれが置かれた環境や個体差によってその見

え方が大きく変わるため、判別器の構築が容易でないためである。

このように明示的に判別基準が与えにくいタスクに対しては通常、データから判別器を求める機械学習の手法が有効とされている。そのためには判別の対象としたい物体が映った多数の画像、ならびにそれぞれに付与された正解ラベルからなる訓練データを多

数用意する必要がある。すなわち大規模画像データベースの構築が一般物体認識システムの実現にあたって重要な要素のひとつとなっている。

従来、画像認識システムの構築にあたっては、専用の画像集合が使用されてきた。すなわち多数の画像に手動でメタデータを付与する必要があった。しかし、精度の高い認識システムを実現するためにはより多くの画像に対してメタデータを用意する必要があり、画像認識システムの構築という目的のためだけに手作業で大規模なデータベースを構築するのには限界があった。

2. 研究の目的

画像中に何が映っているかを自動判定する一般物体認識のための画像ならびに画像特徴量データベースを構築する。

基礎的なデータとして使用するため、World Wide Web から画像の自動収集を行うシステムを実装する。さらに、一般物体認識の目的にとって適切と思われる画像特徴量とその空間配置をメタデータ中に存在する語彙と対応づける手法を開発する。

画像特徴量と語彙の間の条件付き確率として表現されるモデルを使用し、学習手法の妥当性も評価する。

3. 研究の方法

画像と語彙に関連づける情報を得るために、World Wide Web 上で公開されている画像データに着目する。これらの画像の多くは HTML で記述された Web ページ中に埋め込まれており、周辺のテキストは画像に写されている対象（オブジェクト）と関連していることが多い。実際、多くの Web 画像検索エンジンはこのような周辺テキストを用いる

ことで、任意のキーワードを用いた画像検索を可能にしている。

本研究ではこれらの周辺テキスト中の語彙と画像に写された対象（オブジェクト）に対して機械学習を適用し、画像・画像特徴量・語彙の相互変換システムを実装する。

画像データのソースとして Web 検索を利用することにより、任意のカテゴリを対象とした画像認識システムを短時間／低コストで実現することが可能になる。

さらに本研究の特色として、画像とそれに映っている物体を表すラベルという対応付けだけでなく、元の画像における局所的な範囲から構成される中間的な特徴量のデータベースを構築し、物体における配置に関する確率的な関係を保存することが挙げられる。これによってデータベースは階層的に構成される。

実際、物体は複数の中間的な特徴量の組み合わせによって定義されていると考えられる。生物や人間が物体を認識する際も、画素の輝度値から物体を判別する関数を一括して学習するのではなく、輝度値の分布を中間的な特徴量に変換する関数、および中間的な特徴量の組み合わせに基づき物体の判別を行う関数を別個に学習していると考えられる。このような階層的な認識のために最適な中間的な特徴量のデータベースを構築することで、より精度の高い物体認識システムを実現する。

本手法はモジュラーな形で設計されているため、中間的な特徴量の獲得手法、ならびに物体との対応付けに関しては個別に改善させていくことができる。特に後者に関しては現在使用している確率モデルに基づくものの以外の学習手法と組み合わせて使用することも可能である。

4. 研究成果

World Wide Web から画像と語彙の組を収集し、その間の対応関係を蓄積するシステムを構築した。このデータを用いて画像集合中からその単語において表されるもっとも典型的な画像の候補を求める手法を提案し、実験によって評価した。

さらに、画像特徴量とその空間配置の一般的な分布を学習させるシステムを開発した。これらの手法によって一般物体認識のためのプロトタイプシステムを構築することができた。

さらに、得られた中間的な画像特徴量を用いて、画像を典型性によってランキングするシステムを開発した。本研究では典型性をキーワードに対する画像の適切性を表す指標として定義している。同じラベルを与えられた画像の中でも、人間が判断した際、より典型的とみなす画像と、例外的とみなす画像がある。多くの場合、ユーザが求めるのは典型的な画像である。そこで本研究によって構築された中間的な特徴量を利用し、それらをより顕著に含む画像ほど典型的とみなし、ランキングを行う手法を開発した。これによってユーザの検索クエリをラベルに持つ画像集合中より、もっとも典型的な画像を見つけ、提示を行うシステムを実装した。この手法は現在の画像検索エンジンにおいて起こりがちな、無関係な画像が多数検索結果に表れるという状況を改善させる効果を持つと期待される。

これらの結果に関して、学術雑誌 *Journal of Networks* における原著論文 “Image Retrieval with Generative Model for Typicality”、情報処理学会論文誌（トランザクション）データベースにおける原著論文「アスペクトモデルを用いた Web 画像検索結果からの典型的画像の取得」、ならびに国

際会議 The 21st International Conference on Database and Expert Systems Applications (DEXA 2010) における発表 “Typicality Ranking of Images Using the Aspect Model” という形で発表を行った。

さらに、本研究で使用された確率モデルであるアスペクトモデルをライフログに対する索引付けに利用し、精度向上を行う研究に取り組んだ。これによってアスペクトモデルはライフログにおけるコンテキストの推定に使用でき、無関係な情報のフィルタリングに使用できることが明らかになった。この結果を国際会議 The 16th International Conference on Database Systems for Advanced Applications (DASFAA) において、Lifelog Search System Using a Topic Model for Reducing Recognition Errors という題目で発表した。

今後、さらなる精度向上を実現させるため、確率モデルとは異なる観点を導入するものとして、スパース符号化を用いた画像認識モデルの研究を開始している。この研究では情報を高次元ベクトル空間上の疎なベクトルとして表す符号化を行うというものであり、中間的な特徴量が物体を構成する構成要素に対応していることを考慮した場合、現在までの手法の自然な拡張とみなせるものである。このような発展をはじめとして、本研究における成果を多様な形で今後につなげていく予定である。

5. 主な発表論文等

（研究代表者、研究分担者及び連携研究者には下線）

〔雑誌論文〕（計 2 件）

1. Taro Tezuka and Akira Maeda, Image Retrieval with Generative Model for

Typicality, Journal of Networks, Vol.6,
No.3, pp.387-399, March 2011.(査読有)

2. 手塚太郎, 前田亮, アスペクトモデルを用いた Web 画像検索結果からの典型的画像の取得, 情報処理学会論文誌 (トランザクション) データベース, Vol. 3, No. 2, pp. 13-26, 2010 年 6 月. (査読有)

[学会発表] (計 2 件)

1. Taro Tezuka and Akira Maeda, Audio Lifelog Search System Using a Topic Model for Reducing Recognition Errors, The 16th International Conference on Database Systems for Advanced Applications (DASFAA 2011), The Chinese University of Hong Kong, Hong Kong, China, April 24th, 2011.
2. Taro Tezuka and Akira Maeda, Typicality Ranking of Images Using the Aspect Model, The 21st International Conference on Database and Expert Systems Applications (DEXA 2010), University of Deusto, Bilbao, Spain, September 2nd, 2010.

6. 研究組織

(1) 研究代表者

手塚 太郎 (TEZUKA TARO)

筑波大学・図書館情報メディア系・准教授

研究者番号: 40423016